

Testing the homogeneity of the Universe using gamma-ray bursts

Ming-Hua Li¹ * and Hai-Nan Lin² **

¹ Institute of Astronomy and Space Science, Sun Yat-Sen University, Guangzhou 510275, China

² Department of Physics, Chongqing University, Chongqing 401331, China

September 11, 2015

ABSTRACT

Aims. The discovery of a statistically significant clustering in the distribution of gamma-ray bursts (GRBs) has recently been reported. Given that the cluster has a characteristic size of 2,000–3,000 Mpc and a redshift between $1.6 \leq z \leq 2.1$, it has been claimed that this structure is incompatible with the cosmological principle of homogeneity and isotropy of our Universe. In this paper, we study the homogeneity of the GRB distribution using a subsample of the Greiner GRB catalogue, which contains 314 objects with redshift $0 < z < 2.5$ (244 of them discovered by the Swift GRB Mission). We try to reconcile the dilemma between the new observations and the current theory of structure formation and growth.

Methods. To test the results against the possible biases in redshift determination and the incompleteness of the Greiner sample, we also apply our analysis to the 244 GRBs discovered by Swift and the subsample presented by the Swift Gamma-Ray Burst Host Galaxy Legacy Survey (SHOALS). The real space two-point correlation function (2PCF) of GRBs, $\xi(r)$, is calculated using a Landy-Szalay estimator. We perform a standard least- χ^2 fit to the measured 2PCFs of GRBs. We use the best-fit 2PCF to deduce a recently defined homogeneity scale. The homogeneity scale, R_H , is defined as the comoving radius of the sphere inside which the number of GRBs $N(< r)$ is proportional to r^3 within 1%, or equivalently above which the correlation dimension of the sample D_2 is within 1% of $D_2 = 3$.

Results. For a flat Λ CDM Universe, a best-fit power law, $\xi(r) = (r/r_0)^{-\gamma}$, with the correlation length $r_0 = 413.64 \pm 135.40 h^{-1}\text{Mpc}$ and slope $\gamma = 1.57 \pm 0.63$ (at 1σ confidence level) for the real-space correlation function $\xi(r)$ is obtained. We obtain a homogeneous distribution of GRBs with correlation dimension above $D_2 = 2.97$ on scales of $r \geq 8,200 h^{-1}\text{Mpc}$. For the Swift subsample of 244 GRBs, the correlation length and slope are $r_0 = 387.51 \pm 132.75 h^{-1}\text{Mpc}$ and $\gamma = 1.57 \pm 0.65$ (at 1σ confidence level). The corresponding scale for a homogeneous distribution of GRBs is $r \geq 7,700 h^{-1}\text{Mpc}$. For the 75 SHOALS GRBs, the results are $r_0 = 288.13 \pm 192.85 h^{-1}\text{Mpc}$ and $\gamma = 1.27 \pm 0.54$ (at 1σ confidence level), with the homogeneity scale $r \geq 8,300 h^{-1}\text{Mpc}$. For the 113 SHOALS GRBs at $0 < z < 6.3$, the results are $r_0 = 489.66 \pm 260.90 h^{-1}\text{Mpc}$ and $\gamma = 1.67 \pm 1.07$ (at 1σ confidence level), with the homogeneity scale $r \geq 8,700 h^{-1}\text{Mpc}$.

Conclusions. The results help to alleviate the tension between the new discovery of the excess clustering of GRBs and the cosmological principle of large-scale homogeneity. It implies that very massive structures in the relatively local Universe do not necessarily violate the cosmological principle and could conceivably be present.

Key words. Gamma-ray burst: general – Methods: data analysis – Methods: statistical – Cosmology: large-scale structure of Universe – Cosmology: observations – Cosmology: distance scale

1. Introduction

Gamma-ray bursts (GRBs) are the most energetic events in our Universe. Their cosmological origin has been studied by Klebesadel et al. (1973); Meegan et al. (1992); Kouveliotou et al. (1993); Costa et al. (1997); Paradijs et al. (1997); Harrison et al. (1999); Meszaros & Gehrels (2012). GRBs and luminous red galaxies (LRGs) are both luminous tracers of matter in our Universe. Unlike most LRGs, GRBs have a larger redshift range that reaches up to $z \sim 8$. They have minimum separations over 100 Mpc. Therefore, they are valid indicators of potential large-scale structures in the intermediate-redshift Universe.

According to modern cosmology, structures on different length scales in our Universe all have their origins in the quantum fluctuations of the inflation field, the scalar field which generates the inflation after the birth of our Universe. These primordial Gaussian random phase fluctuations, which later lead to the density fluctuations in different modes, enter the horizon

on different epochs and grow as time passes by. This gives rise to the hierarchical scenario of the matter clustering – the fluctuations with longer comoving wavelengths enter the horizon later and have less time to evolve than the small-scale fluctuations. As large-scale structures are closely related to these long-wavelength modes, one can conclude that a finite age of our Universe would result in a limited maximum size of the large-scale structure in our Universe. Translated into the language of modern cosmology, it is the cosmological principle – the matter distribution in our Universe is homogeneous and isotropic on sufficiently large scales. On small scales, our Universe is inhomogeneous, with structures like galaxies and galaxies clusters.

The transition scale between homogeneity and inhomogeneity is called the ‘homogeneity scale’¹. The homogeneity scale of the distribution of matter has long been studied and the results are quite scattered. Hogg et al. (2005) analysed the enormous LRG sample of the Sloan Digital Sky Survey (SDSS) (York et al. 2000) and presented a homogeneity scale $R_H \sim 70$

* limh@ihep.ac.cn

** linhn@ihep.ac.cn

¹ Alternatively, one can call it the inhomogeneity scale. Here we use the conventional term that was used in Yadav et al. (2010).

h^{-1} Mpc. Similar results were obtained by Sarkar et al. (2009) and Scrimgeour et al. (2012), who performed a multifractal analysis over 200,000 blue luminous galaxies in the WiggleZ survey (Drinkwater et al. 2010). Labini et al. (2009) claimed to find a homogeneity scale above $100 h^{-1}$ Mpc after studying the galaxy catalogue of SDSS.

Horvath et al. (2014) have recently reported the discovery of a statistically significant clustering in the GRB sample between $1.6 \leq z \leq 2.1$. They called it the Hercules–Corona Borealis Great Wall (Her-CrB GW). It has a characteristic scale of $\sim 2,000$ Mpc and its longest dimension $\sim 3,000$ Mpc, is six times larger than the size of the Sloan Great Wall. Its characteristic size is far above the upper limit of the homogeneity scale placed by the fractal dimensional analysis based on the galaxy surveys. The two-dimensional Kolmogorov-Smirnov test (Lopes et al. 2008) shows a 3σ deviation. The clustering excess cannot be entirely attributed to the known sampling biases. The existence of the potential structure, defined by the GRBs, is considered to be inconsistent with the cosmological principle and beyond the standard excursion set theory for structure growth.

In this paper, we examine the spatial distribution of the GRBs using a subsample of Greiner’s GRB catalogue, which contains 314 objects with redshift $0 < z < 2.5$ (244 of them discovered by the Swift GRB Mission). The sample encompasses the redshift region of the reported potential structure mapped by the GRBs. We first use the Landy-Szalay estimator (Landy & Szalay 1993) to estimate the real space two-point correlation function (2PCF) $\xi(r)$ of the GRB sample. We fit a simple power law to the measured GRB $\xi(r)$. We then use the best-fit $\xi(r)$ to deduce the correlation dimension $D_2(r)$ and the homogeneity scale R_H of the GRB distribution. To test the results against the possible biases in redshift determination and the incompleteness of the Greiner sample, we also apply the analysis to the subsample presented by the Swift Gamma-Ray Burst Host Galaxy Legacy Survey (SHOALS) in Perley et al. (2015). The results are plotted in the same figure for comparison.

The rest of the paper is organized as follows. In Section 2, we introduce the GRB sample and the techniques used in our analysis. The real-space 2PCF $\xi(r)$ of the GRB sample and its best-fit power law are both obtained in this section. In Section 3, the definitions and calculations of homogeneity scale and correlation dimension are presented. We deduce the specific R_H for the GRBs with $0 < z < 2.5$. Relevant physical implications and comparisons to the results of other surveys are presented in Section 4.

2. Data and techniques

2.1. GRB catalogue and the subsample

In this work we primarily use a sample of 314 GRBs with redshift $0 < z < 2.5$. All of these GRBs are from the collection presented by J. Greiner (2014)²; 244 of them come from the NASA Swift Mission; and the rest come from BeppoSAX GRBM, HETE2, IPN, and INTEGRAL, etc. We use the data released on July 8, 2015. The entire sample contains more than 1,000 objects. Only 431 of them have well-measured redshifts. The well-measured subsample has a redshift $0 < z < 9.2$. In the redshift region $z > 7$, there are only two GRBs: GRB 090423 ($z = 8.26$) and GRB 090429B ($z = 9.2$). They are omitted from our numerical analysis since they are of little statistical significance. This leaves a subsample of 429 GRBs at $0 < z < 6.7$.

Given that the potential GRB structure reported by Horvath et al. (2014) has a redshift $1.6 < z < 2.1$, we cut this sample down further by using only those GRBs having well-determined redshifts at $0 < z < 2.5$, which encompasses the redshift region of the potential structure. We have a final subsample of 314 GRBs at redshift $0 < z < 2.5$.

The redshifts of GRBs can be estimated in a number of ways, i.e. through the absorption spectroscopy of the optical afterglow (the vast majority of the Swift sample are measured in this way) or measuring the emission lines of their host galaxies (which is observationally expensive). Faintness of the afterglows of some events have left the sample intrinsically optically biased. Alternative approaches have been proposed to solve this problem (Kruhl et al. 2011; Rossi et al. 2012; Perley et al. 2013; Hunt et al. 2014). One of them is to develop a series of observability cuts with a set of optimized parameters to isolate a subset of the GRB sample (Jakobsson et al. 2006; Cenko et al. 2006; Perley et al. 2009; Greiner et al. 2011). Using this method, one can obtain a GRB sample whose afterglow redshift completeness is close to about 90% (Hjorth et al. 2012; Jakobsson et al. 2012; Kruhl et al. 2012; Salvaterra et al. 2012; Schulze et al. 2015), the level that is necessary for systematic biases not to dominate the statistical ones (Perley et al. 2015). Using this technique, Perley et al. (2015) established the largest and most complete (92% completeness) GRB redshift sample to date, which is called the SHOALS sample. The SHOALS sample contains 119 objects in total, 112 of which have well-determined redshifts at $0 < z < 6.3$ (75 are at $0 < z < 2.5$). To test the results against the possible biases in redshift determination and the incompleteness of the Greiner sample, we also apply our analysis to this subsample of GRBs.

Thus, we base our analysis primarily on the Greiner sample of 314 GRBs. In comparison, we also apply our analysis to the Swift subsample of 244 GRBs and the SHOALS very complete GRB sample. The angular and redshift distributions of these samples are plotted in Figure 1a and 1b. The corresponding celestial coordinates and the redshifts of the objects in the Greiner sample and the SHOALS sample are listed respectively in Tables 1 and 2, which are publicly available online³. Most of them have a comoving separation $r > 100 h^{-1}$ Mpc, with the longest separation distance up to $r \sim 10 h^{-1}$ Gpc.

2.2. Two-point correlation function

Given a GRB in the spatial volume dV_1 , the 2PCF of GRBs, $\xi(r)$, is defined as the probability of finding another GRB in dV_2 with a separation distance r , i.e. (Peebles 1980)

$$dP_{12} = \bar{n}^2 [1 + \xi(r)] dV_1 dV_2, \quad (1)$$

where \bar{n} is the mean number density of the GRBs. To calculate $\xi(r)$, an auxiliary random sample of N_R points is generated in a window W of observations. A window W is a three-dimensional space of volume V , the same volume as that on which the observation was made.

A statistical estimation of $\xi(r)$ involves a pair count of neighbouring GRBs at a given separation scale. The most widely used estimator of the 2PCF is the Landy-Szalay estimator $\hat{\xi}_{LS}(r)$ (Landy & Szalay 1993),

$$\hat{\xi}_{LS}(r) = \frac{DD(r) - 2DR(r) + RR(r)}{RR(r)}, \quad (2)$$

³ Tables 1 and 2 are only available in electronic form at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via <http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/>

² <http://www.mpe.mpg.de/~jcg/grbgen.html>

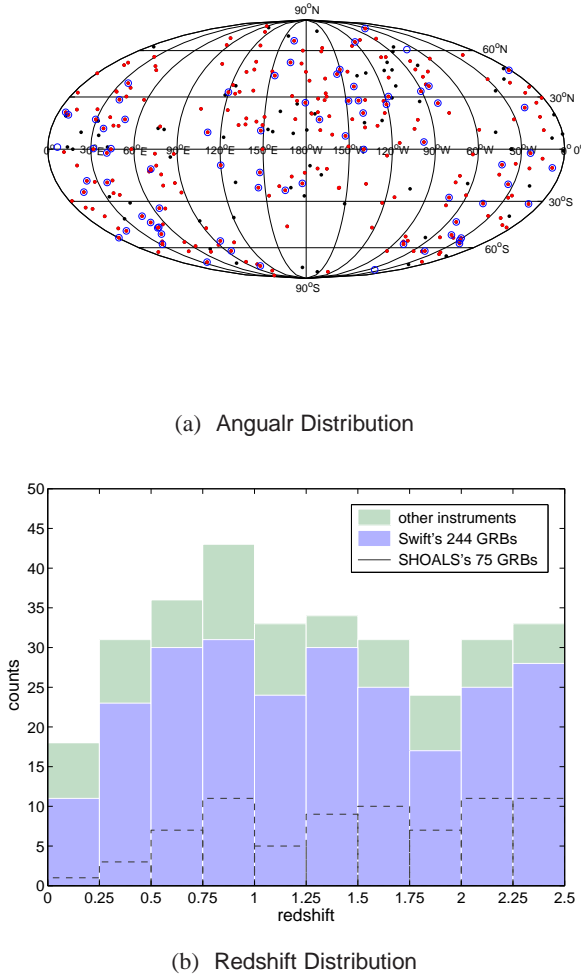


Fig. 1. Angular and redshift distributions of the GRB samples. (a) The angular distribution of the GRB samples in J.2000 equatorial coordinates. The red solid dots represent the 244 GRBs at $0 < z < 2.5$ detected by Swift, while the black solid dots represent those discovered by other detectors within the same redshift range. The red and black solid dots constitute Greiner's GRB sample of 314 objects at $0 < z < 2.5$. The blue circles represent the 112 GRBs (at $0 < z < 6.3$) from SHOALS. (b) The redshift distribution of the GRB data. The y -axis denotes the number of objects in each redshift bin. The green shaded area plus the purple area indicates the total of 314 GRBs from the Greiner sample. The dashed line represents the contribution from the SHOALS subsample of 75 objects at $0 < z < 2.5$.

where $DD(r)$ and $RR(r)$ are, respectively, the number of GRB pairs within the separation $d \in [r - \Delta r/2, r + \Delta r/2]$ (Δr is the bin width used in the statistical estimation of $\xi(r)$) in the observed data set D and in the auxiliary random sample R in the window W , while $DR(r)$ is the number of GRB pairs between the observed data and the random sample with the same separation. The parameter d is the comoving separation distance of GRBs. Specifically, $DR(r)$ is defined as $DR(r) \equiv N_{DR}(r)/(N_D N_R)$ (N_D and N_R are, respectively, the total number of GRBs in the data set D and in the random sample R), where $N_{DR}(r)$ is given by (Kerscher et al. 2000)

$$N_{DR}(r) = \sum_{\mathbf{x} \in D} \sum_{\mathbf{y} \in R} F(\mathbf{x}, \mathbf{y}). \quad (3)$$

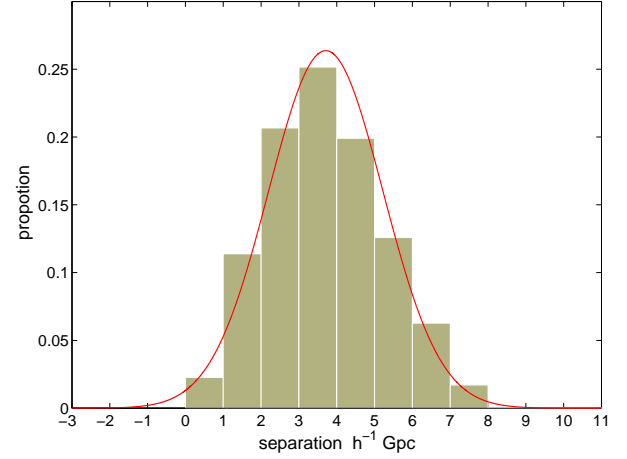


Fig. 2. The distribution of the comoving separations d between the GRBs in the Greiner sample (including the Swift subsample of 244 GRBs and most of the SHOALS GRBs). The x -axis is the comoving separation of the GRB data, in units of h^{-1} Gpc. The y -axis is the proportion of the GRB number in each bin to the whole sample, which has been normalized to 1. The separation d obeys a Gaussian distribution with the expectation value $\bar{d} \approx 3.710 h^{-1}$ Mpc.

The summation runs over all the coordinates of GRBs (represented by \mathbf{x} and \mathbf{y}) in the observed data set D and the random sample R in the window W . The value of the function $F(\mathbf{x}, \mathbf{y})$ equals 1 when the separation of the two objects is within the distance $d(\mathbf{x}, \mathbf{y}) \in [r - \Delta r/2, r + \Delta r/2]$ or otherwise equals 0. The expressions $DD(r) \equiv N_{DD}(r)/[N_D(N_D - 1)]$ and $RR(r) \equiv N_{RR}(r)/[N_R(N_R - 1)]$ respectively represent the normalized number of GRB pairs within the separation mentioned above in the observed data set D and in the random sample R ; $N_{DD}(r)$ and $N_{RR}(r)$ are defined in a similar way as $N_{DR}(r)$ in equation (3).

We use the jackknife resampling method to determine the statistical uncertainty of the measured 2PCF of GRBs. The jackknife method is an internal method of error estimation that is extensively used to determine the errors of 2PCF of galaxies and quasars (Ross et al. 2007; Sawangwit et al. 2011; Nikoloudakis et al. 2013). The entire sample is divided into N' subsamples of roughly equal size. The jackknife error estimator is given as

$$\sigma_{\text{Jack}}^2(r) = \sum_{i'=1}^{N'} \frac{DR_{i'}(r)}{DR(r)} [\xi_{i'}(r) - \xi(r)]^2, \quad (4)$$

where $\xi_{i'}(r)$ denotes the estimate of the 2PCF on all of the $(N' - 1)$ subsamples except the i -th one.

2.3. Calculating the two-point correlation function, $\xi(r)$

The estimation of the 2PCF of GRBs, $\xi(r)$, is made by counting the pairs in and between the observed GRB distribution and a catalogue of randomly distributed GRBs. The density of random points that we use for the estimation is 20 times the density of the GRB data. We use a bin width of $\Delta \log_{10}(r) = 0.2 h^{-1}$ Mpc. The estimation of $\xi(r)$ depends on the assumed cosmology. We use a flat Λ CDM cosmological model (cold dark matter plus a cosmological constant Λ) in this work, with $\Omega_\Lambda = 0.72$, $\Omega_m = 0.28$, $H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$, $h = 0.7$.

Furthermore, the 2PCF is usually measured in redshift space. As shown in Figure 2, the GRBs in the observed sample have an

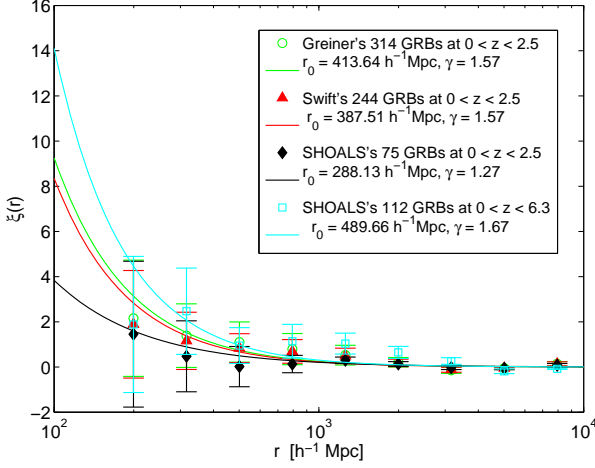


Fig. 3. Best-fit power law of the measured real-space 2PCF $\xi(r)$ at $200h^{-1} < r < 10^4 h^{-1}$ Mpc. We assume a flat Λ CDM cosmological model, with $\Omega_\Lambda = 0.72$, $\Omega_m = 0.28$, $H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$, $h = 0.7$. The real-space 2PCF $\xi(r)$ measured via equation (2) for the Greiner, Swift, and SHOALS samples are, respectively, indicated by green circles, red solid triangles, black solid diamonds, and cyan squares with 1σ jackknife error bars that are estimated from (4). The density of random points we use for the estimation is 20 times the density of GRB data. $\xi(r)$ is plotted in equally spaced logarithmic intervals of $\Delta \log_{10}(r) = 0.2 h^{-1} \text{ Mpc}$. The best-fit power laws of the form (5) for the measured $\xi(r)$ are plotted in solid lines, with the best-fit parameters given in the legends.

average separation distance over $100 h^{-1} \text{ Mpc}$ (with the largest separation up to $\sim 10 h^{-1} \text{ Gpc}$). On such a scale, the current structure formation theory predicts that the evolution and clustering of matter should still be in the linear regime at the present time (Springel et al. 2005; Eisenstein et al. 2007). The redshift-space distortions⁴ due to the small-scale peculiar velocities of the objects and the redshift variances are also minimal on this scale (Ross et al. 2007). Thus, the difference between the redshift-space and the real-space correlation functions on such large scales could be negligible. For convenience, the calculation and analysis in this paper are done in real space.

For the estimation of the jackknife error $\sigma_{\text{jack}}^2(r)$, we take $N' = 5$, and split the sample into five redshift regions with equal redshift intervals $\Delta z = 0.5$. The real-space 2PCFs in a flat Λ CDM Universe for Greiner's 314 GRBs, the subsample of 244 objects discovered by Swift, and the SHOALS subsample are respectively plotted in Figure 3 in equally spaced logarithmic intervals. It spans a comoving distance of scale $10^2 h^{-1} < r < 10^4 h^{-1}$ Mpc.

A power law of the form

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma} \quad (5)$$

is usually fitted to the correlation functions of galaxies and galaxy clusters (Davis & Peebles 1983; Bahcall 1988; Maddox et al. 1990; Peacock & West 1992; Dalton et al. 1994; Zehavi et al. 2004). The parameter r_0 is the comoving correlation length, in units of $h^{-1} \text{ Mpc}$. The slope γ is a dimensionless constant. Similarly, we fit a power law of the form in Eq. (5) to the measured 2PCF data over the range $200h^{-1} \leq r \leq 10^4 h^{-1}$

⁴ A more formal description of 'redshift distortions' can be found in Section 9.4 in Dodelson (2008).

Mpc. We perform a standard least- χ^2 fit. For a flat Λ CDM Universe, the best-fit values of the correlation length r_0 and slope γ for the real-space correlation function $\xi(r)$ for Greiner's 314 GRB data at $z < 2.5$ are (with 1σ confidence level errors)

$$r_0 = (413.64 \pm 135.40) h^{-1} \text{ Mpc}, \quad (6)$$

$$\gamma = 1.57 \pm 0.63, \quad (7)$$

with $\chi_{\text{min}}^2 = 0.73t$. For the subsample of 244 objects discovered by Swift, we obtain (with 1σ confidence level errors)

$$r_0 = (387.51 \pm 132.75) h^{-1} \text{ Mpc}, \quad (8)$$

$$\gamma = 1.57 \pm 0.65, \quad (9)$$

with $\chi_{\text{min}}^2 = 0.90$.

For the SHOALS subsample of 75 GRBs, the results are

$$r_0 = (288.13 \pm 192.85) h^{-1} \text{ Mpc}, \quad (10)$$

$$\gamma = 1.27 \pm 0.54, \quad (11)$$

with $\chi_{\text{min}}^2 = 0.50$. For the SHOALS sample of 113 GRBs at $0 < z < 6.3$, the results are

$$r_0 = (489.66 \pm 260.90) h^{-1} \text{ Mpc}, \quad (12)$$

$$\gamma = 1.67 \pm 1.07, \quad (13)$$

with $\chi_{\text{min}}^2 = 2.82$. From these results, one can see that the best-fit values of r_0 and γ increase with the growing number of GRB data points as well as higher redshifts. We plot the best-fit power-law models of $\xi(r)$ in Figure 3.

3. Homogeneity of the GRB distribution

Given the best-fit 2PCF, $\xi(r)$, for the GRB sample in the previous section, we are now in a position to calculate the homogeneity scale, R_H , for the GRB distribution at $0 < z < 2.5$. We first give a brief introduction of the correlation dimension $D_2(r)$ for a random distribution of data points. We describe the relation between the value of $D_2(r)$ and the concept of a homogeneous distribution. A more formal treatment of this section can be found in Yadav et al. (2010) and Scrimgeour et al. (2012).

3.1. Correlation dimension, $D_2(r)$

Several methods have been developed to investigate the homogeneity of the galaxy distribution. The most popular among them is fractal analysis (Yadav et al. 2005). A fractal is a kind of geometrical object where every small part of it appears as a reduction of the entirety. In fractal analysis, the concept 'fractal dimension' is invoked to describe the homogeneity of the distribution of a point set. One of the most common definitions of fractal dimension is the 'correlation dimension', $D_2(r)$. Unlike other homogeneity indicators, deviations caused by a size-limited sample would only result in second-order changes to $D_2(r)$. Thus, $D_2(r)$ is regarded as a robust measure of homogeneity and is extensively used in the homogeneity investigations of galaxies and quasars (Labini et al. 2009; Scrimgeour et al. 2012; Nadathur 2013).

In this paper, we use the working definition of $D_2(r)$ given in Scrimgeour et al. (2012) to study the homogeneity of the GRB distribution. We limit our discussions to a three-dimensional

space. Given a set of points in space, the measurement of $D_2(r)$ is to find the average number of neighbouring points, $N(< r)$ inside a three-dimensional sphere of radius r centered at each point. One can formulate the scaling behaviour of $N(< r)$ as

$$N(< r) \propto r^D, \quad (14)$$

where D is the fractal dimension of the distribution. $N(< r)$ can actually be given as $N(< r) = 4\pi r^3 \bar{n}/3$, where \bar{n} is the mean number density of points in that region. For a homogeneous distribution of the point set, \bar{n} is a universal constant and $N(< r)$ then scales as $\propto r^3$. From equation (14), this implies $D = 3$. An inhomogeneous distribution would result in a non-universal \bar{n} . From equation (14), this gives $D < 3$ or $D > 3$: $D < 3$ stands for an inhomogeneous distribution of the point set, while $D > 3$ represents a ‘super-homogeneous’ distribution (Gabrielli et al. 2002). In the literature, $N(< r)$ is usually divided by the number expected for a homogeneous distribution, $4\pi r^3 \bar{n}/3$ (here \bar{n} is a universal constant), to correct for incompleteness. The corrected $N(< r)$ is given as

$$N(< r) \propto r^{D-3}. \quad (15)$$

For a homogeneous distribution, one has $N(< r) \propto r^{3-3} = 1$.

In general, the correlation dimension D in equation (14) is a function of the sphere radius, r , i.e. $D = D_2(r)$. It can be deduced from the count-in-sphere number $N(< r)$ as

$$D_2(r) \equiv \frac{d \ln N(< r)}{d \ln r} = \frac{d \ln N(< r)}{d \ln r} + 3. \quad (16)$$

The correlation dimension $D_2(r)$ measures the scaling properties of $N(< r)$ without being affected by the amplitude of $N(< r)$ (which is related to the mean number density of the data points in the regions.) It is therefore an objective measure of the homogeneity in the statistical analysis of matter distribution. To deduce the correlation dimension $D_2(r)$ for the GRB distribution, one should first calculate $N(< r)$. For the GRBs $N(< r)$ can actually be obtained by integrating the 2PCF $\xi(r)$ for the GRBs (Peebles 1980):

$$N(< r) = \frac{3}{4\pi r^3} \int_0^r [1 + \xi(r')] 4\pi r'^2 dr'. \quad (17)$$

Combining equation (16) with (17), we obtain an analytical expression of $D_2(r)$:

$$D_2(r) = \frac{d}{d \ln r} \left\{ \ln \left[\frac{3}{4\pi r^3} \int_0^r [1 + \xi(r')] 4\pi r'^2 dr' \right] \right\} + 3. \quad (18)$$

Given the best-fit power law of $\xi(r)$ as in equation (5) with the correlation length r_0 and slope γ of $\xi(r')$ given in equations (6) to (13), we are now in a position to estimate the homogeneity scale of the GRB distribution.

3.2. Scale of homogeneity, R_H

There are several ways to define the homogeneity scale R_H with respect to the correlation dimension $D_2(r)$. Yadav et al. (2010) defined the homogeneity scale as the scale above which the deviation of the fractal dimension $D_2(r)$ from the ambient spatial dimension becomes smaller than the statistical dispersion of $D_2(r)$ itself. For the limited size of the GRB sample that we use in our work, the statistical dispersion of $D_2(r)$ might be large. This effect of a limited sized sample would therefore enlarge the estimation of R_H for our GRB sample. Following Scrimgeour et al. (2012), in our analysis we use a more robust definition of R_H that

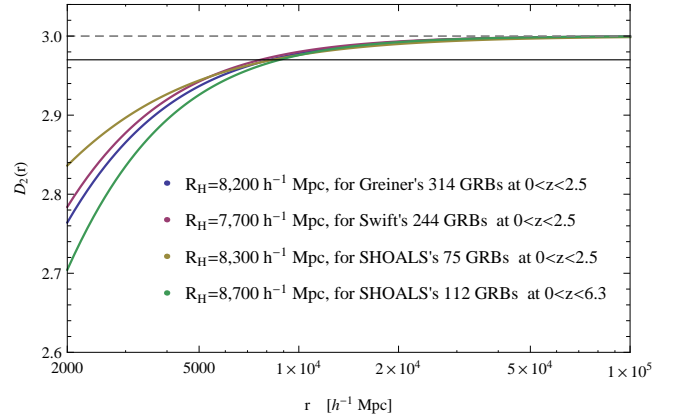


Fig. 4. Correlation dimension, $D_2(r)$, for different GRB samples. The $D_2(r)$ calculated from equation (18) is shown in solid lines. The dashed line indicates the critical value defined for the transition from a homogeneous to an inhomogeneous distribution of the GRBs, i.e. 1% from the homogeneity, $D_2(r) = 2.97$.

is not affected by the sample size. Given the correlation dimension $D_2(r)$ of the GRB distribution, the homogeneity scale R_H is defined as the scale on which $D_2(r)$ of the sample is within 1% of $D_2 = 3$, i.e. $D_2(r = R_H) = 2.97$. An equivalent definition of R_H can also be given by $N(r)$, for which R_H is defined as the co-moving radius r of the sphere inside which the number of LQGs $N(< r)$ is proportional to r^3 within 1%, i.e. $N(r = R_H) = 1.01$.

Using these definitions, we calculate the homogeneity scale of the distribution of the GRB sample. The theoretical prediction of $D_2(r)$ and the corresponding value of R_H for the GRB samples are plotted in Figure 4 for comparison. For Greiner’s sample of 314 GRBs, the homogeneity scale for the GRB distribution is $R_H \simeq 8,200 h^{-1} \text{Mpc}$, which means that the cosmological principle is retained on such a scale ($r > R_H$). One should observe a homogeneous distribution of GRBs on scales $r > R_H$. For the subsample of 244 GRBs discovered by Swift, the result is $R_H \simeq 7,700 h^{-1} \text{Mpc}$. For the SHOALS subsample of 75 GRBs at $0 < z < 2.5$, we have $R_H = 8,300 h^{-1} \text{Mpc}$. For the SHOALS subsample of 113 GRBs at $0 < z < 6.3$, we have $R_H = 8,700 h^{-1} \text{Mpc}$.

The scale of homogeneity R_H for the GRBs can also be considered as the upper limit of the characteristic size of any clustering and structures in the GRB distribution. The potential structure mapped by the GRBs at $1.6 < z < 2.1$ reported by Horvath et al. (2014) has a characteristic size of $\sim 2,000 \text{ Mpc}$ and is thus well within the limits we obtain. Therefore, we conclude that the existence of such an excess clustering in the GRB distribution is still compatible with the cosmological principle and the standard theory of structure formation.

4. Conclusions and discussion

Redshift surveys (Drinkwater et al. 2010; Ahn et al. 2014) provide about a hundred thousand galaxies that can be used for a homogeneity investigation of the matter distribution in our Universe. Most of these galaxies have redshifts $z < 1$. GRBs usually have a larger redshift range than the galaxies (which can reach up to $z \sim 8$) and thus provide valid luminous indicators of matter distribution in the intermediate-redshift universe. Quasars have redshifts as high as GRBs, but their observations

only cover a limited sky area⁵. This involves specific technical treatment in data process when using the Landy-Szalay estimator to estimate the 2PCF for quasars (Karagiannis et al. 2014). Compared with galaxies and quasars, the GRB sample has a full-sky angular distribution and a redshift range up to $z \sim 8$. In fact, in our study we found that most GRBs have a comoving separation $> 100 h^{-1}\text{Mpc}$ (see Figure 2) and therefore they provided an efficient probe of matter correlation on large scales in the intermediate-redshift universe. By ‘efficient’ here, we mean that the size of the data we used to estimate $\xi(r)$ on such a scale is much less than that used by galaxy or quasar surveys, since the number density of galaxies and quasars observed is much higher than GRBs. Using galaxies or quasars to probe the correlation and clustering on such a scale (e.g. $r \sim 200 h^{-1}\text{Mpc}$) would involve counting many more objects inside the sphere of radius r and therefore would cost more computing machine time than using GRBs.

In this paper we used a sample of 314 GRBs from the collection presented by J. Greiner (2014) to study the homogeneity of matter distribution on large scales. They cover a redshift range $0 < z < 2.5$, 244 of which discovered by the Swift GRB Mission. We calculated the real-space 2PCF, $\xi(r)$, for GRBs on scales $10^2 h^{-1} < z < 10^4 h^{-1} \text{Mpc}$. The measured $\xi(r)$ for GRBs on scales $r > 200 h^{-1}\text{Mpc}$ can be well fit by a power law of the form $\xi(r) = (r/r_0)^{-\gamma}$, with correlation length $r_0 = (413.64 \pm 135.40) h^{-1}\text{Mpc}$ and slope $\gamma = 1.57 \pm 0.63$ (1σ confidence level), and $\chi^2_{\min} = 0.73$. For the subsample of 244 objects discovered by Swift, we obtained $r_0 = (387.51 \pm 132.75) h^{-1}\text{Mpc}$ and $\gamma = 1.57 \pm 0.65$ (1σ confidence level), with $\chi^2_{\min} = 0.90$.

We then used the best-fit $\xi(r)$ to deduce the homogeneity scale R_H for the GRB distribution at $0 < z < 2.5$. We obtained $R_H \approx 8,200 h^{-1}\text{Mpc}$. For the subsample of 244 objects discovered by Swift, we obtained $R_H \approx 7,700 h^{-1}\text{Mpc}$, which means that above such a scale GRBs can be considered to have a homogeneous distribution. On scales $r < R_H$, an inhomogeneous distribution of GRBs is assumed for the standard excursion set theory of structure growth. The potential GRB structure recently reported by Horvath et al. (2014) has a characteristic size of $\sim 2,000 \text{Mpc}$ with its longest dimension $\sim 3,000 \text{Mpc}$. Both are well below the homogeneity scale R_H we deduced for the GRB distribution. This implies that the discovery of such an angular excess of the GRB distribution is still compatible with the cosmological principle, which assumes that the matter distribution of our Universe is homogeneous and isotropic over a large smoothing scale. For the distribution of GRBs, we suggested that such a scale is $R_H \approx 8,200 h^{-1}\text{Mpc}$ (which corresponds to $R_H \approx 11.7 \text{Gpc}$ given the choice of $h = 0.7$). The comoving cosmic horizon⁶ within which causality holds are $l_H \approx 10 \text{Gpc}$ at $z \approx 2$ (the redshift of the potential GRB structure). The R_H we obtained for the GRB distribution is slightly larger than the l_H at $z \approx 2$. We hope that with the growing size of the observed GRB sample and a more accurate measurement of the 2PCF $\xi(r)$ of GRBs, this difference will be eliminated by future observations and investigations.

To test the results against the possible biases in redshift determination and the incompleteness of the Greiner sample, we also applied our analysis to the GRB subsample presented by SHOALS in Perley et al. (2015). Despite its limited sample size (with a total of 119 objects), the afterglow redshift completeness

of the SHOALS GRB sample is 92%. It was proposed in order to strike a balance between redshift completeness and overall statistical size; the SHOALS GRB sample provides largest and most complete GRB redshift sample for clustering analysis to date. It has been employed to test the viability of our results against the biases and incompleteness of the Swift and Greiner samples. We obtained $r_0 = 288.13 \pm 192.85 h^{-1}\text{Mpc}$ and $\gamma = 1.27 \pm 0.54$ (at 1σ C.L.), with $\chi^2_{\min} = 0.50$ for the SHOALS 75 GRBs at $0 < z < 2.5$. The deduced homogeneity scale is $R_H = 8,300 h^{-1}\text{Mpc}$. For the 113 objects in the SHOALS sample that have well-determined redshifts at $0 < z < 6.3$, we obtained $r_0 = 489.66 \pm 260.90 h^{-1}\text{Mpc}$ and $\gamma = 1.67 \pm 1.07$ (at 1σ C.L.), with $\chi^2_{\min} = 2.82$. The deduced homogeneity scale is $R_H = 8,700 h^{-1}\text{Mpc}$. Two comments are necessary. First, the best-fit values of r_0 and γ increase with growing number of GRB data points as well as higher redshifts. This manifests itself in the results obtained from SHOALS’s GRB sample, and also in the results obtained from the Greiner sample. Second, one can see that the R_H obtained from the SHOALS very complete GRB sample do not differ much from those obtained from the Greiner and Swift subsets, which implies that the Greiner GRB sample and the Swift GRB sample are both robust in the statistical analysis of clustering and matter distribution.

It should also be noted that for distributions of different objects, the scale of homogeneity R_H are different because the different objects have different origins and evolution dynamics, and thus have different bias factors. The 2PCF can also be given by $\xi(r) \equiv b^2 \xi_{\text{mass}}(r)$, where b is the bias factor for specific kinds of objects. Given the same 2PCF $\xi_{\text{mass}}(r)$ of the underlying (dark) matter⁷, the amplitude of $\xi(r)$ depends on the value of b . In general, b is a function of comoving radius r and redshift z , i.e. $b = b(r, z)$. Thus, for different kinds of objects with different characteristic (mass) scales and redshifts, $b(r, z)$ take different values, which would result in a varying amplitude of $\xi(r)$. From equation (18), one can expect a different homogeneity scale R_H for the distribution of different objects. For galaxies, R_H takes a value of $\sim 70\text{--}100 h^{-1}\text{Mpc}$ (Yadav et al. 2005; Hogg et al. 2005; Sarkar et al. 2009; Labini 2011; Scrimgeour et al. 2012). For quasars in the DR7QSO catalogue (Schneider et al. 2010), the value is $R_H \sim 180 h^{-1}\text{Mpc}$ (Nadathur 2013). For GRBs with redshift $0 < z < 2.5$, we obtained a $R_H \approx 8,600 h^{-1}\text{Mpc}$. Our study provides a supplement to the clustering analysis based on the galaxies (Yadav et al. 2005; Hogg et al. 2005; Sarkar et al. 2009; Labini 2011; Scrimgeour et al. 2012) and quasars (White et al. 2012; Nadathur 2013; Karagiannis et al. 2014).

Homogeneity on large scales is one of the cornerstones of modern cosmological theory. The large-scale homogeneity in the very early Universe (at redshift $z \approx 1100$) is well supported by the high degree of isotropy of the cosmic microwave background radiation power spectrum (Bennett et al. 2013) and the PLANCK data (Ade et al. 2013). Although recently the discoveries of several large quasar groups (LQGs) and structures have been reported, e.g. the CCLQG (i.e. U1.28) and U1.11 (Clowes et al. 2012), the Huge-LQG (Clowes et al. 2013), and the Her-CrB GW (Horvath et al. 2014), the method used for assessing the statistical significance and overdensity of the groups varies with individuals. Nadathur (2013) provided a fractal dimension analysis of the DR7 quasar catalogue and found that the quasar distribution is homogeneous above the scale $\sim 130 h^{-1}\text{Mpc}$. Changbom et al. (2012) carried out a large cos-

⁵ The quasar catalogue from the SDSS DR10 contains 166,583 quasars detected over $6,373 \text{deg}^2$. Over half of them have redshifts $z > 2.15$.

⁶ The comoving cosmic horizon is given by $l_H = \int_0^a da' / H(a')$, where $H(a)$ is the Hubble factor and a is the cosmic scale factor.

⁷ The underlying matter 2PCF, $\xi_{\text{mass}}(r)$ is given by the Fourier transform of the primordial matter power spectrum, $P(k)$ which is generated by the cosmic inflation (Scrimgeour et al. 2012).

mological N -body simulation to demonstrate that the existence of the Sloan Great Wall and a void complex in the SDSS region is perfectly consistent with the Λ CDM model. Considering all these results, it should be prudent for one to claim that the recent discoveries actually contradict the cosmological principle of homogeneity. In fact, Nadathur (2013) suggested that the homogeneity scale is an average property. It is not necessarily affected by the discovery of a single large structure. It implies that very massive structures in the relatively local Universe could conceivably be present.

Despite the scattering results, objects like quasars and LQGs and events like GRBs provide new tools to study the matter distribution and any possible power excess in the intermediate- and high-redshift Universe. We hope that the next generation of sky surveys will offer excellent prospects for clearing up the perplexities between the observations of large-scale structures and the standard excursion set theory of structure formation.

Acknowledgements. This work was supported by the National Natural Science Foundation of China (No. 30000-41030521). We warmly thank Prof. Zhi-Bing Li from the School of Physics and Engineering, Sun Yat-Sen University, and Prof. Miao Li from the Institute of Astronomy and Space Science, Sun Yat-Sen University, for enlightening discussions and comments. This work is based on the GRB catalogue presented by Jochen Greiner at <http://www.mpe.mpg.de/~jcg/grbgen.html> and the Swift Gamma-Ray Burst Host Galaxy Legacy Survey (“SHOALS”) in Perley et al. (2015) (arXiv:astro-ph1503.04246). We would also like to thank the anonymous editor and referee for their informative comments and constructive suggestions in improving the manuscript.

References

- Ade, P. A. R., Aghanim, N., Armitage-Caplan, C. et al. [Planck Collaboration] 2013, *A&A*, 1303, 5062.
- Ahn, C., Hee Kim, S., Gu Yun, M. et al. 2014, *ApJS*, 211, 17.
- Bahcall, N. A. 1988, *ARA&A*, 26, 631.
- Bennett, C. L., Larson, D., Weiland, J. L. et al. 2013, *ApJS*, 208, 20.
- Cenko, S. B., et al. 2006, *PASP*, 118, 1396
- Changbom, P., Yun-Young, C., Juhan, K. et al. 2012, *ApJL*, 759, L7.
- Clowes, R. G., Campusano, L. E., Graham, M. J. et al. 2012, *MNRAS*, 419, 556.
- Clowes, R. G., Harris, K. A., Raghunathan, S. et al. 2013, *MNRAS*, 429, 2910.
- Costa, E., Frontera, F., Heise, J. et al., 1997, *Nature*, 387, 783
- Dalton, G. B., Efstathiou, G., Maddox, S. J. et al. 1992, *ApJL*, 390, L1.
- Dalton, G. B., Croft, R. A. C., Efstathiou, G. et al. 1994, *MNRAS*, 271, L47.
- Davis, M. & Peebles, P. J. E. 1983, *ApJ*, 267, 465.
- Dodelson, S., 2008, *Modern Cosmology*, Elsevier(Singapore) Pte Ltd.
- Drinkwater, M. J., Jurek, R. J., Blake, C. et al. 2010, *MNRAS*, 401, 1429.
- Eisenstein, D. J., Seo, Hee-Jong; Sirko, E. et al. 2007, *ApJ*, 664, 675.
- Gabrielli, A., Joyce, M., & Labini, F. 2002, *Phys. Rev. D*, 65, 083523.
- Greiner, J., Krühler, T., Klose, S. et al. 2011, *A&A*, 526, A30
- Harrison, F., Bloom, J. S., Frail, D. A. et al., *ApJL*, 523, L121
- Hermit, S., Santiago, B. X., Lahav, O. et al. 1996, *MNRAS*, 283, 709.
- Horváth, I., Hakkila, J., & Bagoly, Z. 2014, *A&A*, 561, L12.
- Hjorth, J., Malesani, D., Jakobsson, P. et al. 2012, *ApJ*, 756, 187.
- Hogg, D. W., Eisenstein, D. J., Blanton, M. R. et al. 2005, *ApJ*, 624, 54.
- Hunt, L. K., Palazzi, E., Michalowski, M. J. et al. 2014, *A&A*, 565, A112
- Jakobsson, P., Hjorth, J., Malesani, D. et al. 2012, *ApJ*, 752, 62.
- Jakobsson, P., Levan, A., Fynbo, J. P. U. et al. 2006, *A&A*, 447, 897
- Karagiannis, D., Shanks, T., Ross, N. P. 2014, *MNRAS*, 441, 486.
- Kerscher, M., Szapudi, I. & Szalay, A. 2000, *ApJL*, 535, L13.
- Klebesadel, R., Strong, I., & Olson, R. 1973, *ApJL*, 182, L85
- Kouveliotou, C., Meegan, C. A., Fishman, G. J. et al. 1993, *ApJ*, 413, 101
- Krühler, T., Greiner, J., Schady, P. et al. 2011, *A&A*, 534, A108
- Krühler, T., Malesani, D., Milvang-Jensen, B. et al. 2012, *ApJ*, 758, 46
- Labini, F. S., Vasilyev, N. L., Pietronero, L. et al 2009, *Europhys. Lett.*, 86, 4900.
- Labini, F. S. 2011, *Europhys. Lett.*, 96, 59001.
- Lopes, R. H. C., Hobson, P. R., & Reid, I. D. 2008, *J. Phys. Conf. Ser.*, 119, 042019.
- Landy, S. D. & Szalay, A. 1993, *ApJ*, 412, 64.
- Maddox, S. J., Efstathiou, G., Sutherland, W. J. et al. 1990, *MNRAS*, 242, 43P.
- Meegan, C., Fishman, G. J., Wilson, R. B. et al. 1992, *Nature*, 355, 143
- Mészáros, P. & Gehrels, N. 2012, *Res. Astron. Astrophys.*, 12, 1139.
- Nadathur, S. 2013, *MNRAS*, 414, 398.
- Nikoloudakis N., Shanks T., & Sawangwit U. 2013, *MNRAS*, 429, 2032.
- Nichol, R. C., Collins, C. A., Guzzo, L. et al. 1992, *MNRAS*, 255, 21p.
- van Paradijs, J., Groot, P. J., Galama, T. et al. 1997, *Nature*, 386, 686
- Peacock, J. A. & West, M. J. 1992, *MNRAS*, 259, 494.
- Peebles, P. J. E. 1980, *The Large-Scale Structure of the Universe*. Princeton Univ. Press, Princeton, NJ.
- Perley, D. A., Cenko, S. B., Bloom, J. S. et al. 2009, *AJ*, 138, 1690
- Perley, D. A., Levan, A. J., Tanvir, N. R. et al. 2013, *ApJ*, 778, 128
- Perley, D. A., Krühler, T., Schulze, S. et al. 2015, submitted to *ApJ*, arXiv:astro-ph/1504.02482v2.
- Postman, M., Huchra, J. P., & Geller, M. J. 1992, *ApJ*, 384, 404.
- Ross, P. N., DaAngela, J., Shanks, T. et al. 2007, *MNRAS*, 381, 573.
- Rossi, A., Klose, S., Ferrero, P. et al. 2012, *A&A*, 545, A77
- Salvaterra, R., Campana, S., Vergani, S. D. et al. 2012, *ApJ*, 749, 68.
- Sarkar, P., Yadav, J., Pandey, B. et al. 2009, *MNRAS*, 399, L128.
- Sawangwit, U., Shanks, T., Croom, S. M. et al. 2011, *MNRAS*, 416, 3033
- Schneider, D. P., Richards, G. T., Hall, P. B. et al. 2010, *AJ*, 139, 2360
- Schulze, S., Chapman, R., Hjorth, J. et al. 2015, *ApJ*, 808, 73.
- Scrimgeour, M., Davis, T., Blake, C. et al. 2012, *MNRAS*, 425, 116.
- Springel, V., White, S. D. M., Jenkins, A. al. 2005, *Nature*, 435, 629.
- White, M., Myers, A. D., Ross, N. P. et al. 2012, *MNRAS*, 424, 933.
- York, D., Adelmann, J., Anderson, J. E. et al. 2000, *AJ*, 120, 1579.
- Yadav, J., Bharadwaj, S., Pandey, B. et al. 2005, *MNRAS*, 364, 601.
- Yadav, J., Bagla, J., & Khandai, N. 2010, *MNRAS*, 405, 2009.
- Zehavi, I., Blanton, M. R., Frieman, J. A. et al. 2002, *ApJ*, 571, 172.
- Zehavi, I., Weinberg, D. H., Zheng, Z. et al. 2004, *ApJ*, 608, 16.